Unlocking European marine biodiversity under EMODnet Biology data using the FAIR principles

Simon Claus, Vlaams Instituut voor de Zee (Belgium), simon.claus@vliz.be
Christos Arvanitidis, Hellenic Centre for Marine Research (Greece), arvanitidis@her.hcmr.gr
Nicoals Bailly, Hellenic Centre for Marine Research (Greece), nbailly@hcmr.gr
Klaas Deneudt, Vlaams Instituut voor de Zee (Belgium), Klaas.deneudt@vliz.be
Daphnis De Pooter, Vlaams Instituut voor de Zee (Belgium), daphnis.depooter@vliz.be
Peter Herman, Deltares (The Netherlands), Peter.Herman@deltares.nl
Dan Lear, Marine Biological Association (United Kingdom), dble@MBA.ac.uk
Paula Oset, Vlaams Instituut voor de Zee (Belgium), paula.oset.garcia@vliz.be
Leen Vandepitte, Vlaams Instituut voor de Zee (Belgium), leen.vandeitte@vliz.be
EMODnet Biology Partnership (http://www.emodnet-biology.eu/partners)

Marine biodiversity data are essential to measure and study the ecosystem health of maritime basins. These data are often collected with limited spatial and temporal scope and are scattered over different organizations in small datasets for a specific species group or habitat. Therefore there is a continuous need to assemble these individual datasets, and process them into interoperable biological data products. The European Marine Observation and Data Network (EMODnet), supported by the EU's integrated maritime policy aims therefore to provide a single access point to European marine biodiversity data and products by assembling individual datasets from various sources and process them into interoperable data products. EMODnet Biology is a long term marine data initiative and is structuring its activities around four main pillars:

Making Marine Biological Data Findable

Detailed inventories of the recent and historical European biological datasets are created and published online via the EMODnet biology data catalogue. This catalogue is based on IMIS, an ISO-19115 compliant metadata system and contains over 1,200 dataset descriptions with information on the what, who, when, where and why a dataset was collected. These datasets include many large monitoring data collections from different European regional seas representing observations of marine species of phytoplankton, zooplankton, macro-algae, angiosperms, benthos, birds, mammals, reptiles or fish. For several of them a Digital Object Identifier (DOI) is equally available.

Making Marine Biological Data Accessible

Today, already 875 EMODnet Biology datasets, representing over 23 milion occurrence records are freely accessible through the EMODnet Biology data download portal. The portal includes 1. a toolbox allowing federated selection queries over different datasets, 2. a description on how the

Fig. 1 - EMODnet-Biology data portal at www.emodnet-biology.eu.

API can be accessed (as a WFS service) if a user wants to access machine readable marine biology data and 3. a link to the IPT resources, being the raw data available in Darwin Core Archive. The data are integrated into the European Ocean Biogeographic Informaiton System (EurOBIS) datasystem which contributes to OBIS, an global strategic alliance of people and organizations sharing a vision to make marine biogeographic data, from all over the world, freely available over the World Wide Web. There are now three subsections under data download.

Making Marine Biological Data Interoperable

All marine biological data is transformed to a common data structure and passes taxonomic (using the WORMS Vocabulary) and geographic (using the Marineregions Vocabulary) quality control procedures. The common data structure and standards used within EMODnet Biology consists of a DarwinCore (Dwc) Event Core in combination with a DwC Occurrence Extension and an enhancement to the DwC MeasurementOrFact Extension. This new structure enables the linkage of measurements or facts - quantitative and qualitative properties - to both sampling events and species occurrences, and includes additional fields for property standardization (using the NERC Vocabulary). The standard also allows to organize, aggregate, and link ocean observation events using “event hierarchy”.

Making Marine Biological Data Reusable

Through liaison with key stakeholders including governments, regulatory authorities, academia, NGOs and industry, it is clear that the real value lies in the the development of scalable, information-rich data products, based on high-quality underlying data. The products developed within EMODnet Biology are now being structured around the Essential Ocean Variables for biodiversity and will be provided through an Atlas of Marine Life. They inlcude 1. Trait based analysis, 2. Gridded species abundance maps, 3. Time series analysis and 4. Species distribution models.